2015 European Control Conference (ECC)
July 15-17, 2015. Linz, Austria

A numerical evaluation of state reconstruction methods for heterogeneous cell populations

Steffen Waldherr Robert Frysch Tim Pfeiffer Theresa Jakuszeit Shen Zeng Georg Rose

Abstract— **Heterogeneity among cells is a common characteristic of living systems. For mathematical modeling of heterogeneous cell populations, one typically has to reconstruct the underlying heterogeneity from measurements on the population level. Based on recent insights into the mathematical nature of this problem as an inverse problem of tomographic type, we evaluate numerical methods to perform such a reconstruction in basic case studies. We compare a kernel density based optimization approach, filtered back projection, and algebraic reconstruction techniques. The latter two are well established methods in computed tomography.**

I. Introduction

Mathematical models of heterogeneous cell populations have been studied at least since the 1960s [2]. These models are based on formulating the heterogeneity probabilistically as a density function. However, while the modeling framework was theoretically formulated for arbitrary high-dimensional intracellular networks, in practice commonly only one or two intracellular variables were used, which were also available for measurement [11]. For such models, state estimation would not have been required. In recent years, researchers are also integrating population heterogeneity in much more complex models of intracellular signaling [6]. In these situations, not all intracellular variables can be measured, and state estimation for population models becomes important for the modeling. Typical experimental data is in the form of population snapshots from high-throughput single cell measurements, and can be described mathematically as density functions using statistical density estimation methods [14]. Ad hoc optimization methods have been used since some years for state estimation [15], [4], [17], but a systems theoretic analysis of identifiability properties has only been performed recently [16], [18].

A first necessary condition for identifiability of heterogeneous population models from population snapshot data was presented in [16], based on a novel measure theoretical approach to formulate such models. In subsequent contributions, a sufficient and necessary condition for identifiability was discovered, which is based on viewing identifiability as a tomographic reconstruction problem [18], [19]. In computed tomography, one considers the problem of reconstructing a density function from its integrals over affine subspaces [7], [12], which is exactly the problem one is facing when reconstructing a state density from measured output densities.

The objective of the study reported in this paper is to evaluate numerical methods for reconstructing the state density from output densities in heterogeneous cell population models. Specifically, we compare optimization-based methods based on kernel density approximations [4] and two methods from computed tomography [12] with each other in numerical case studies. In addition, we evaluate the effect of quantitative observability of the underlying single cell model on the reconstruction quality of the state density.

*This work was partly funded by the German federal state Sachsen-Anhalt through the Center for Dynamical Systems – Biosystems Engineering, by the German Federal Ministry of Education and Research (BMBF) within the Forschungscampus STIMULATE under grant number 03FO16101A, and by the state of Saxony-Anhalt (grant number I 60).

SW and TJ are with the Institute for Automation Engineering, RF, TP, and GR are with the Institute for Medical Engineering, Otto-von-Guericke-University, 39106 Magdeburg, Germany. SZ is with the Institute for Systems Theory and Automatic Control, University of Stuttgart, 70550 Stuttgart, Germany. Correspondence to steffen.waldherr@ovgu.de

II. The state reconstruction problem in heterogeneous cell populations

A. Problem definition

In this section, we recall the problem definition for state reconstruction in heterogeneous cell populations, as introduced in a measure theoretical setting by [16], [18].

For this paper, we focus on linear cell models, as the tomographic methods have so far only been transferred to linear systems [18]. The model for the dynamics of an individual cell is given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where the state vector $x \in \mathbb{R}^n$ may consist of both intracellular variables and constant parameters, the latter with trivial dynamics. The heterogeneity among individuals is described by the initial condition $x(0) = x_0$, which is a random vector subject to the probability distribution

$$x_0 \sim \mathbb{P}_0. \tag{2}$$

The heterogeneity in the initial condition leads to a heterogeneity in the output $y(t) \in \mathbb{R}^m$, which can thus be considered as a random vector within the population. The output is characterized by the time-varying probability distribution $\mathbb{P}_{y(t)}$. This distribution is given as the push-forward distribution of the initial distribution $\mathbb{P}_0$ under the system dynamics (1). Based on a Borel algebra $\mathcal{B}(\mathbb{R}^m)$, the push-forward distribution is computed as the pull-back measure [16], [18]

$$\mathbb{P}_{y(t)|\mathbb{P}_0}(B_y) = \mathbb{P}_0((Ce^{At})^{-1}(B_y)) \tag{3}$$

for a measurable set $B_y \in \mathcal{B}(\mathbb{R}^m)$, where $(Ce^{At})^{-1}$ denotes the preimage of Ce^{At}.

The following definition of identifiability for the population model (3) has been proposed.

Definition 1 ([16]): A population model is called *structurally identifiable*, if the following implication holds:

$$\big(\forall\, t \geq 0 :\, \mathbb{P}_{y(t)|\mathbb{P}_0'} = \mathbb{P}_{y(t)|\mathbb{P}_0''}\big) \quad \Rightarrow \quad \mathbb{P}_0' = \mathbb{P}_0'', \tag{4}$$

where $\mathbb{P}_0'$ and $\mathbb{P}_0''$ are arbitrary probability distributions.
A sufficient condition for a population model to be structurally identifiable is that the individual cell model (1) is observable, and that $\operatorname{rank} C = n - 1$ [18].

In this paper, we consider numerical approaches to solve the reconstruction problem for a finite number of measured output distributions: Given the measured probability distributions $\mathbb{P}_{y(t_k)}$ for $k = 1, \ldots, N$, determine the initial distribution $\mathbb{P}_0$. We aim at models which satisfy the sufficient condition $\operatorname{rank} C = n-1$. This condition appears to be tight for generic observable systems (A, C).

Generally, the numerical approaches discussed in this paper are based on representing the probability distributions $\mathbb{P}_{y(t)}$ and $\mathbb{P}_0$ with density functions $p_{y(t)} : \mathbb{R}^m \to \mathbb{R}$ and $p_0 : \mathbb{R}^n \to \mathbb{R}$, respectively. To simplify notation, define the matrix $\Phi_t = \begin{pmatrix} C \\ C^\perp \end{pmatrix} e^{At}$, where $C^\perp$ is a matrix that spans the kernel of C, with $\operatorname{rank} C^\perp = 1$ due to $\operatorname{rank} C = n - 1$. From (3), the density functions p_0 and $p_{y(t)}$ are related by the following integral over an affine subspace [18]:

$$p_{y(t)}(y) = |\det \Phi_t^{-1}| \int_{z \in \mathbb{R}} p_0\big(\Phi_t^{-1} \begin{pmatrix} y \\ z \end{pmatrix}\big) \mathrm{d}z. \tag{5}$$

The reconstruction problem then amounts to reconstructing the initial density function p_0 from measurements of the output density functions $p_{y(t_k)}$ at the time instances $t_1, \ldots, t_N$.

B. Optimization based reconstruction method

The optimization based reconstruction method is based on parametrizing the unknown initial density as is commonly done in non-parametric density estimation [14]. In this approach, the initial density is approximated with a finite set of basis functions $k : \mathbb{R}^n \to \mathbb{R}$, centered at grid points from the grid

$$\mathcal{G} = \{x_G^{(1)}, \ldots, x_G^{(M)}\}, \tag{6}$$

where M is the number of grid points. A common choice for the basis function k is the Gaussian kernel

$$k(x) = (2\pi)^{-n/2} e^{-x^{\mathrm{T}} x/2}. \tag{7}$$

The approximation of the initial density is constructed as

$$p_0(x) \approx \sum_{i=1}^{M} \alpha_i \frac{k\big((x - x_G^{(i)})/h\big)}{h^n}, \tag{8}$$

where $h > 0$ is the window width for the kernels, and $\alpha_i \in \mathbb{R}$ are the weight coefficients to be estimated in the optimization problem. In practice, the window width h should be chosen in relation to the distance between grid points: from our experience, half the grid width is a reasonable window width. Substituting (8) into the push-forward density (5), we obtain

$$p_{y(t)}(y) \approx \sum_{i=1}^{M} \alpha_i \int_{z \in \mathbb{R}} \frac{k\Big(\big(\Phi_t^{-1} \begin{pmatrix} y \\ z \end{pmatrix} - x_G^{(i)}\big)/h\Big)}{h^n |\det \Phi_t|} \mathrm{d}z. \tag{9}$$

The integral over the kernel function k can be computed *a priori* without any measurement data. Using the resulting coefficient densities

$$\psi_{i,t}(y) = \int_{z \in \mathbb{R}} \frac{k\Big(\big(\Phi_t^{-1} \begin{pmatrix} y \\ z \end{pmatrix} - x_G^{(i)}\big)/h\Big)}{h^n |\det \Phi_t|} \mathrm{d}z \tag{10}$$

for $i = 1, \ldots, M$, we can formulate the problem of reconstructing the kernel approximation to the initial density as a least squares optimization problem. Let $\hat{p}_{y(t_j)}$ be the measured output density at time t_j, with $j = 1, \ldots, N$ running over the number N of available time points. The optimization problem is then given as

$$\begin{aligned} &\min_{\alpha_i \geq 0} \sum_{j=1}^{N} \|\hat{p}_{y(t_j)} - \sum_{i=1}^{M} \alpha_i \psi_{i,t_j}\|^2 \\ &\text{s.t. } \sum_{i=1}^{M} \alpha_i = 1. \end{aligned} \tag{11}$$

For optimization, it is useful to base the optimization on the L^2 norm of the deviation in density, even though other density-based approaches to dynamical systems focus more on the L^1 norm [10]. A quadratic norm has also been used in earlier studies [4], where a kernel density based optimization problem has been used in a more heuristic, not measure based, approach.

For numerical solution, the density function norm in the optimization problem (11) is approximated via a discretization in the argument space. The most straightforward approximation of the norm is by evaluating the density function over a grid $\mathcal{H}$ and summing up the squared function values, yielding

$$\|p\|^2 \approx \sum_{y \in \mathcal{H}} p(y)^2 \Delta y^2 \tag{12}$$

as approximation for the norm of a density function p. The function values $p(y)$ over the grid $y \in \mathcal{H}$ can be collected in a vector

$$P = \big(p(y)\big)_{y \in \mathcal{H}}. \tag{13}$$

With this notation, the optimization problem (11) is approximated by the quadratic program

$$\begin{aligned} &\min_{\alpha \geq 0} \sum_{j=1}^{N} (\alpha^{\mathrm{T}} \Psi_{t_j}^{\mathrm{T}} \Psi_{t_j} \alpha - 2 \hat{P}_{y(t_j)}^{\mathrm{T}} \Psi_{t_j} \alpha + \hat{P}_{y(t_j)}^{\mathrm{T}} \hat{P}_{y(t_j)}) \\ &\text{s.t. } \sum_{i=1}^{M} \alpha_i = 1, \end{aligned} \tag{14}$$

where $\alpha = (\alpha_1, \ldots, \alpha_M)^{\mathrm{T}}$. In view of the identifiability results discussed in Section II-A, positive definiteness of the matrix $\sum_{j=1}^{N} \Psi_{t_j}^{\mathrm{T}} \Psi_{t_j}$ and thus a unique optimal solution to

the quadratic program can be ascertained from observability of the underlying single cell system (1) together with an appropriate choice of time points t_j where measurements are taken. The resulting estimated initial density is then constructed according to (8), where the optimal solution of the quadratic program (14) is substituted for the coefficients α_i.

C. Tomography based reconstruction methods

The integral over an affine subspace of the initial density p_0 in (5) is an integral over parallel lines for each time point. Therefore, it can be identified with the X-ray transform, resp. the Radon transform for $n = 2$ of this function. In the context of computed tomography (CT) these line integrals are referred to as projections. Consequently, a time point in (5) corresponds to a single projection of the image encoding one specific angle θ_t (between normal vector and x-axis) of a set of integrals of parallel lines. Herein, the line integrals represent the measured extinctions that arise from accumulated attenuation coefficients during the propagation of X-rays through the image.

Thus, a well-known problem of reconstructing slice images in CT is obtained. Hence, we have got access to a comprehensive pool of methods for the reconstruction, which can be separated into analytical and iterative algorithms. For comparison, we use two well-known procedures: the analytical filtered back projection (FBP) and the iterative algebraic reconstruction technique (ART). The FBP is an analytical solution of the CT problem (i.e. inverse Radon transformation) based on the Fourier slice theorem. This theorem connects the two-dimensional Fourier space of an image with one-dimensional Fourier transformations of its projections (Radon transformation). In accordance to the description above, we parameterize the density function $p_{y(t)}$ in dependency on the orthogonal distance $s = s(y, \theta_t)$ of the lines to the origin and the corresponding angle θ_t, i.e. $p_{y(t)}(y) = p_{\theta_t}(s)$. The FBP solution can then be formulated as [13]:

$$p_0(x_1, x_2) = \int_0^{\pi} \tilde{p}_{\theta_t}(x_1 \cos\theta + x_2 \sin\theta)\mathrm{d}\theta. \qquad (15)$$

Here, $\tilde{p}_{\theta_t}(s)$ denotes the processed projections after filtering with a ramp filter. Since this filtering has high pass character, it is highly susceptible to noise. The integration in formula (15) represents a backsmearing process for the available angles, i.e. the filtered projection values are added to the reconstructed image along the direction of the incident rays. Figure 1 illustrates the backsmearing procedure exemplarily for projections of an ideal elliptic object.

Note that (15) is analytically exact only if an infinite number of angles covering the full interval $[0, \pi]$ is used. For practical use, both conditions are often not satisfied. This degrades the FBP to an approximate solution which comes along with image artifacts. Especially angular undersampling is a typical scenario, since experimental conditions usually restrict measurements to a few number of time points (streak artifacts). Nevertheless, FBP is a very fast method to obtain reconstructions and is therefore common in CT.

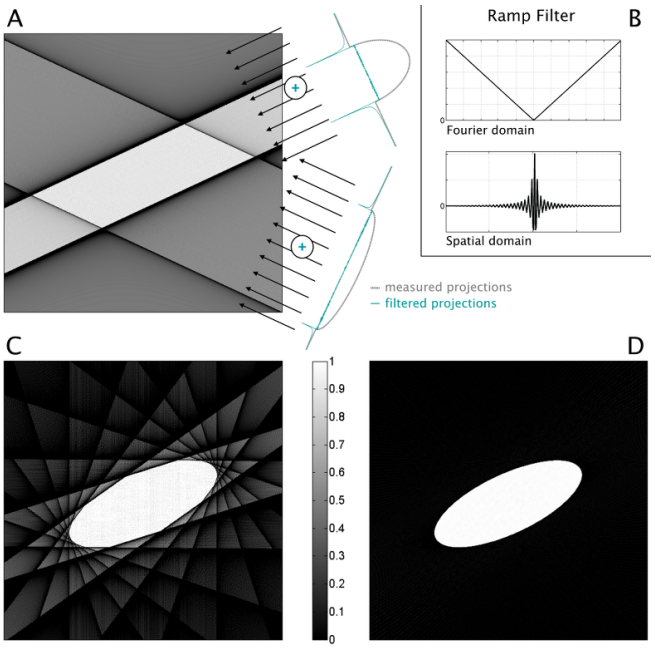

Fig. 1. A) Illustration of backsmearing procedure for filtered projections from two angles: the gray graphs on the right depict measured projections. The result after filtering (cyan) is "smeared back" (backprojected) along the direction indicated by the arrows, i.e. added to the image; B) Fourier and spatial representation of the ramp filter. C) FBP reconstruction using 10 projections and D) 180 projections.

One way to increase the reconstruction quality is to perform an iterative reconstruction. For the ART, problem (5) is reformulated in a vector notation and solved as a linear equation system in terms of finding a minimal solution of

$$\min_{p_0 \geq 0} \left\| S p_0 - \hat{p}_{\{y(t)\}} \right\|. \qquad (16)$$

The matrix S is the discretized formulation of the integration along the lines. A common algorithm to solve problem (16) is the Landweber algorithm [9], which we also use here:

$$p_0^{k+1} = p_0^k - \omega S^{\mathrm{T}}(S p_0^k - \hat{p}_{\{y(t)\}}), \qquad (17)$$

where p_0^k denotes the solution for the initial density function in the k^{th} iteration step and $\omega \in (0, 1]$ is a relaxation parameter. The ART procedure can be described as follows: First, the forward projection of the current reconstruction, denoted by Sp_0^k, is calculated. Afterwards, that result is compared with the measured projections by computation of the difference. The choice of ω then controls to what extent this difference is smeared back (via S^{T}) in order to update the reconstruction. This update can also be interpreted as a gradient descent step. The whole routine is repeated until the norm of the remaining difference drops below a predefined threshold. This tomographic method is very similar to the optimization based approach B using rectangular base functions (pixel grid) instead of Gaussian kernels.

D. An observability measure

Previous studies have argued that quantitative observability properties of the individual cell model (1) influence the precision with which the initial density can be reconstructed. In this paper, we also pursue a numerical study to evaluate this hypothesis. To this end, we define a quantitative observability measure for the individual cell model (1) that is based on the energy contained in the output sequence, and closely related to the empirical observability Gramian [8] for the time points where measurements are taken.

The energy contained in the output sequence for the time points t_j, $j = 1, \ldots, N$, is given by

$$\|(y(t_j))_{j=1}^N\|^2 = x(0)^{\mathrm{T}}\Big(\sum_{j=1}^{N} e^{A^{\mathrm{T}}t_j}C^{\mathrm{T}}Ce^{At_j}\Big)x(0). \qquad (18)$$

Based on this energy, we consider the matrix

$$W = \sum_{j=1}^{N} e^{A^{\mathrm{T}}t_j}C^{\mathrm{T}}Ce^{At_j}. \qquad (19)$$

As an observability measure, we use the inverse condition number of the matrix W [1] given by the ratio of the smallest and largest singular value

$$\frac{\sigma_{min}(W)}{\sigma_{max}(W)}. \qquad (20)$$

This observability measure is applied to compare different parametrization of the example system studied in the next section.

III. Evaluation of reconstruction methods in case studies

A. Models for a numerical case study

We evaluate the numerical reconstruction methods on two underlying single cell systems. The first single cell system is a simple representation of gene expression in cells, modelled by

$$\begin{aligned} \dot{x}_1 &= qx_2 - x_1 \\ \dot{x}_2 &= 0 \\ y &= x_1, \end{aligned} \qquad (21)$$

where x_1 is the concentration of the gene product and x_2 the expression rate which may differ among cells. The parameter $q > 0$ is an overall expression efficiency, for example caused by environmental conditions. This setup assumes that the concentration of the gene product x_1 is measurable on the single cell level, using for example fluorescence measurements.

The model (21) has been studied previously in the context of population identifiability with the parameter $q = 1$ [16], [18]. Here, we are testing different values of q in order to study the influence of the observability measure on the numerical reconstruction results. The resulting observability measures are given in Table I for two sets of measurement time points, the first one being $(0.0, 0.3, 0.9, 1.8)$, and the second one being $(0.0, 0.1, 0.3, 0.6, 0.9, 1.3, 1.8, 2.4, 3.0, 4.0)$. As expected from the system dynamics, observability is significantly impaired for small values of q.

TABLE I

Parameter-dependent observability measures for system (21)

	$q = 0.1$	$q = 0.2$	$q = 0.5$	$q = 1$
4 time points	0.005	0.021	0.125	0.374
10 time points	0.014	0.054	0.300	0.432

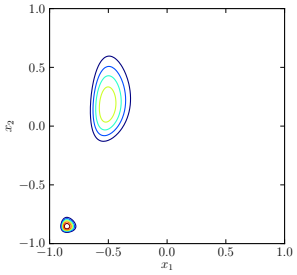

Fig. 2. Contour plots of phantom used for evaluating the reconstruction methods.

The second single cell system is a weakly damped oscillator, defined by

$$\begin{aligned} \dot{x}_1 &= -0.01x_1 - x_2 \\ \dot{x}_2 &= x_1 - 0.01x_2 \\ y &= x_1. \end{aligned} \qquad (22)$$

Oscillations are a relevant dynamic behaviour in biomolecular networks, even though harmonic oscillators are less frequent. From the tomographic point of view, an oscillating system is interesting, since it potentially allows measurements for any desired projection angle.

In tomography, it is common to use so called "phantoms" in order to evaluate both actual machines and algorithms for object reconstruction. Here, we use as phantom a mathematical model for an initital density, given by a bimodal, bivariate log-normal distribution, where the upper mode shows a positive covariance between the two variables (Figure 2). The phantom is modelled after what might be observed in an actual population of living cells, where mixtures of subpopulations are common and individual subpopulations display a distribution similar to a log-normal or Gamma distribution in for example protein expression [5].

Artificial measurement data are generated from the phantoms by solving the forward problem (5) (Figure 3). These data are given to the reconstruction algorithms, and the reconstruction result is compared to the original phantom in order to evaluate the algorithm.

B. Results of numerical reconstruction

To reconstruct the initial density functions from the output densities, we applied the three reconstruction schemes kernel density based optimization, filtered back projection (FBP), and algebraic reconstruction technique (ART) discussed in Sections II-B and II-C.

For the numerical study conducted here, the kernel density based optimization method uses a Gaussian normal kernel

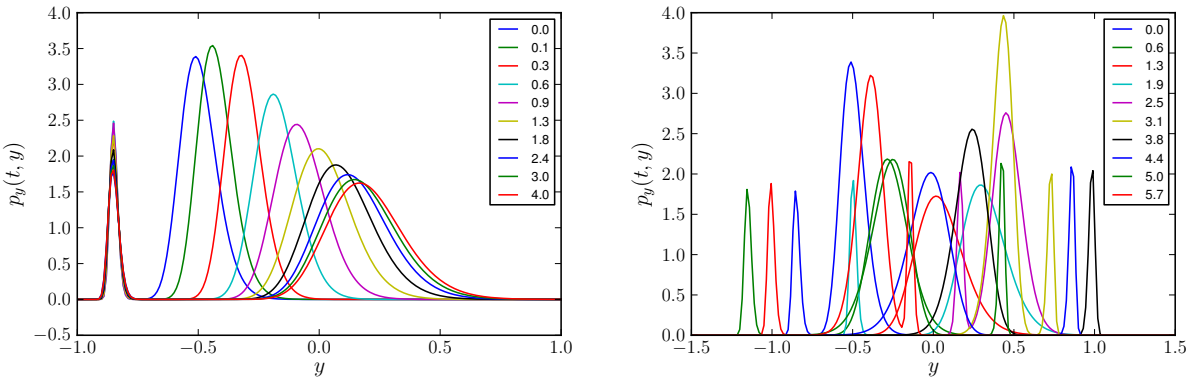

Fig. 3. Output densities from phantom at different time points. Left: System dynamics (21) with $q = 1$. Right: System dynamics (22).

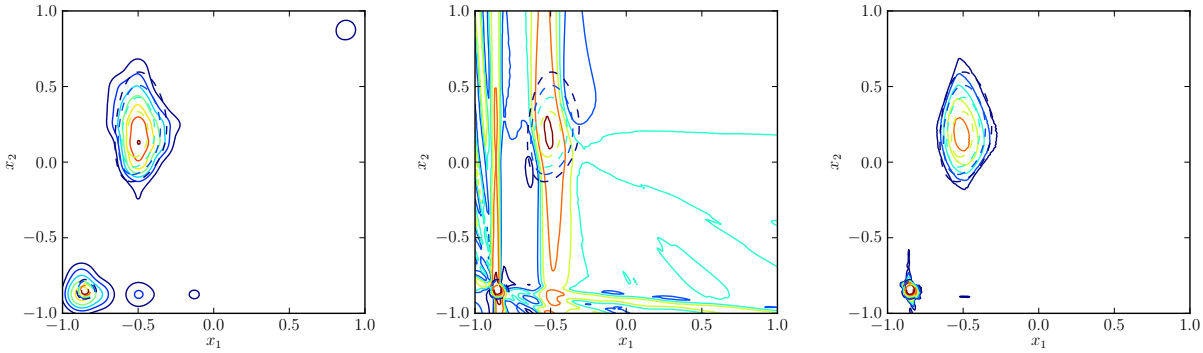

Fig. 4. Reconstruction results from different methods for system (21) with $q = 0.2$. Phantoms are shown as dashed contours, reconstructions as full contours. Left column: Kernel density based reconstruction. Middle column: Tomographic reconstruction (FBP). Right column: Tomographic reconstruction (ART).

with $M = 49$ kernel basis points on a lattice grid and a window width $h = 0.12$ for the data from the normal distribution, and $M = 225$ basis points and $h = 0.06$ for the data from the bimodal distribution. In practice, where the initial density is unknown, the number of basis points and window width could be chosen iteratively, maybe even concentrated in areas where most of the mass of the initial density is estimated to be located. The resulting quadratic program (14) is solved in Python with the cvxopt package[1].

The implementation of FBP and ART reconstruction is done using the pre-implemented routines `radon` and `iradon` from the MATLAB Image Processing Toolbox[2].

The results of the reconstruction are illustrated in Figures 4 and 5. The kernel based optimization and the ART reconstruction yield visually comparable results, and both achieve a reconstruction without or with little artifacts. However, the kernel density based optimization is impaired by the kernel window width, which appears to be not sufficiently narrow to capture the sharp lower mode of the bimodal distribution. The pixel-based ART achieves a very nice reconstruction even of the sharp mode in the bimodal distribution. The FBP reconstruction clearly shows artifacts from the limited number of time points (or projection angles from the tomographic point of view) where measurements were taken.

[1] `http://cvxopt.org`

[2] MATLAB and Image Processing Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States

TABLE II

L^1 ERROR NORMS OF NUMERICAL RECONSTRUCTIONS WITH BIMODAL DISTRIBUTION AS NOMINAL INITIAL DENSITY

System (21)				System (22)
$q = 0.1$	$q = 0.2$	$q = 0.5$	$q = 1$	
Kernel density based optimization				
0.65	0.55	0.53	0.52	0.36
Tomographic FBP				
4.38	3.67	3.15	3.05	7.64
Tomographic ART				
0.22	0.21	0.14	0.13	0.09

For a more quantitative comparison of the reconstruction results, we computed L^1 norms of the difference between the nominal initial density and the reconstruction result (Table II). The significantly weaker performance of the FBP is clearly apparent also from the quantitative comparison. The ART performs better than the kernel density based optimization, in particular for the cases with higher observability measures. Presumably, this is due to the shape and window width of the kernels not being very well chosen for the latter distribution.

For all methods, a lower observability measure in sys-

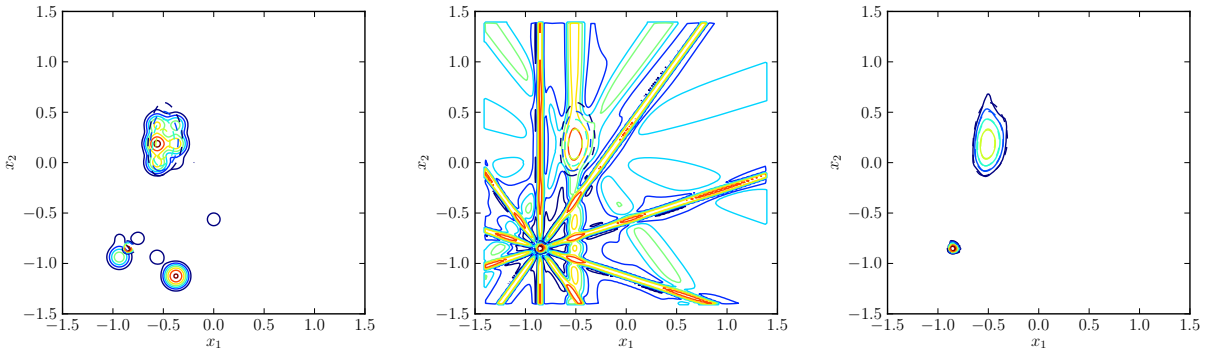

Fig. 5. Reconstruction results from different methods for system (22). Phantoms are shown as dashed contours, reconstructions as full contours. Left column: Kernel density based reconstruction. Middle column: Tomographic reconstruction (FBP). Right column: Tomographic reconstruction (ART).

tem (21) reduced the quality of the reconstruction.

IV. Conclusions

In this paper, we discussed and compared different numerical methods to reconstruct the distribution of the initial condition in population systems from measurements of the output distributions. One method, the kernel density based optimization approach, was used earlier in this context [4]. The other two methods, the filtered back projection (FBP) and the algebraic reconstruction technique (ART) commonly used in computed tomography, were motivated from the recently discovered equivalence of the reconstruction problem to mathematical tomography [18].

The classical tomographic method, the FBP, seems in general not to be well suited for this type of reconstruction problem. This result was expected, since it is already well known that the FBP performs usually badly in situations where only a limited range of projection angles is available [7], [3]. However, the ART seems to offer an advantage over the kernel density based optimization, particularly in cases where a complex initial distribution has to be reconstructed.

Tomographic methods usually focus on reconstruction from integrals over affine subspaces, which would restrict their application to linear models for the intracellular dynamics. Since most intracellular models are nonlinear, it would be of interest to generalize the tomographic approach to integrals over general varieties.

V. Acknowledgements

The authors would like to thank S. Bannasch for QA of mathematical formulations within the manuscript.

References

[1] D. Dochain, N. Tali-Maamar, and J.P. Babary. On modelling, monitoring and control of fixed bed bioreactors. *Computers & Chemical Engineering*, 21(11):1255–1266, 1997.

[2] A.G. Fredrickson, D. Ramkrishna, and H.M. Tsuchiya. Statistics and dynamics of procaryotic cell populations. *Math. Biosci.*, 1(3):327–374, September 1967.

[3] J. Frikel. *Reconstructions in limited angle x-ray tomography: Characterization of classical reconstructions and adapted curvelet sparse regularization*. PhD thesis, Technische Universität München, 2013.

[4] J. Hasenauer, S. Waldherr, M. Doszczak, P. Scheurich, N. Radde, and F. Allgöwer. Analysis of heterogeneous cell populations: A density-based modeling and identification framework. *J. Process Control*, 21(10):1417–1425, 2011.

[5] J. Isensee, M. Diskar, S. Waldherr, R. Buschow, J. Hasenauer, A. Prinz, F. Allgöwer, F. W. Herberg, and T. Hucho. Pain modulators regulate the dynamics of PKA-RII phosphorylation in subgroups of sensory neurons. *J Cell Science*, 127(Pt 1):216–229, Jan 2014.

[6] M. Jeschke, S. Baumgärtner, and S. Legewie. Determinants of cell-to-cell variability in protein kinase signaling. *PLoS Comput Biol*, 9(12):e1003357, 2013.

[7] A. C. Kak and Malcolm Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, 1988.

[8] S. Lall, J. E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Int. J. Robust Nonlin. Control*, 12(6):519–535, May 2002.

[9] L. Landweber. *An Iteration Formula for Fredholm Integral Equations of the First Kind with Application to the Axially Symmetric Potential Flow about Elongated Bodies of Revolution*. University of Maryland, College Park, 1951.

[10] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise: Stochastic aspects of dynamics*. Springer-Verlag, New York, 1994.

[11] J.-J. Liou, F. Srienc, and A. G. Fredrickson. Solutions of population balance models based on a successive generations approach. *Chemical Engineering Science*, 52(9):1529–1540, May 1997.

[12] A. Markoe. *Analytic tomography*, volume 106 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, New York, USA, 2006.

[13] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Ber. Verh. Sächs. Akad. Wiss. Leipzig*, Math. Nat. kl. 69:262–277, 1917.

[14] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

[15] S. Waldherr, J. Hasenauer, and F. Allgöwer. Estimation of biochemical network parameter distributions in cell populations. In *Proc. of the 15th IFAC Symp. Syst. Ident. (SYSID)*, pages 1265–1270, Saint-Malo, France, 2009.

[16] S. Waldherr, S. Zeng, and F. Allgöwer. Identifiability of population models via a measure theoretical approach. In *Preprints of the 19th World Congress of the International Federation of Automatic Control (IFAC)*, pages 1717–1722, Cape Town, South Africa, August 24–29 2014.

[17] C. Zechner, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koeppl. Moment-based inference predicts bimodality in transient gene expression. *Proc Natl Acad Sci U S A*, 109(21):8340–8345, May 2012.

[18] S. Zeng, S. Waldherr, and F. Allgöwer. An inverse problem of tomographic type in population dynamics. In *Proc. 53rd IEEE Conference on Decision and Control (CDC)*, pages 1643–1648, Los Angeles, USA, 2014.

[19] S. Zeng, S. Waldherr, C. Ebenbauer, and F. Allgöwer. Ensemble observability of linear systems. Submitted.